American Economic Review 2016, 106(12): 3765–3799
http://dx.doi.org/10.1257/aer.20151138

Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India†

By Jishnu Das, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan*

We present unique audit-study evidence on health care quality in rural India, and find that most private providers lacked medical qualifications, but completed more checklist items than public providers and recommended correct treatments equally often. Among doctors with public and private practices, all quality metrics were higher in their private clinics. Market prices are positively correlated with checklist completion and correct treatment, but also with unnecessary treatments. However, public sector salaries are uncorrelated with quality. A simple model helps interpret our findings: Where public-sector effort is low, the benefits of higher diagnostic effort among private providers may outweigh costs of potential overtreatment. (*JEL* H42, I11, I18, O15)

Health care is a credence good with substantial information asymmetries between patients and providers. This makes it difficult for patients to determine the quality of care they have received (Dulleck and Kerschbamer 2006). It is widely believed, therefore, that unregulated market-based delivery of health care is socially undesirable. For instance, Arrow (1963, p. 967) notes that "it is the general social consensus, clearly, that the laissez-faire solution for medicine is intolerable." Further,

* Das: The World Bank Group, 1818 H Street NW, Washington, DC 20433 (e-mail: jdas1@worldbank.org); Holla: The World Bank Group, 1818 H Street NW, Washington, DC 20433 (e-mail: aholla@worldbankorg); Mohpal: Department of Economics, University of Michigan, 611 Tappan Avenue, Ann Arbor, MI 48104 (e-mail: amohpal@umich.edu); Muralidharan: Department of Economics, University of California, 9500 Gilman Drive, San Diego, CA 92093 (e-mail: kamurali@ucsd.edu). We thank Brian Chan, Veena Das, Ranendra Kumar Das, Diana Tabak, and Manoj Mohanan for their contribution to designing and implementing the SP methodology, and thank Eun-Young Shim for her contribution to the theoretical Appendix. We thank Prashant Bharadwaj, Gordon Dahl, Roger Gordon, Gordon Hanson, Michael Kremer, Paul Niehaus, and Adam Wagstaff for comments. We are grateful to Innovations for Poverty Action, CT, and Pratap Bhanu Mehta and the Center for Policy Research, New Delhi for hosting the project and providing logistical support and infrastructure. We also thank Sreela Dasgupta, L. Ravi, and Anvesha Khandelwal for project management support, and Monisha Ashok, Carl Liebersohn, Prerna Mukharya, Suzanne Plant, and Anand Shukla for excellent research assistance. The project would not have been possible without the dedication of our SPs and Purshottam, Rajan Singh, Devender, Charu Nanda, Simi Bajaj, Geeta and other staff at the Institute of Socioeconomic Research on Democracy and Development (ISERDD) in Delhi. The Social and Rural Research Institute (SRI), New Delhi oversaw data collection of all other aspects other than SPs. This study was funded by the Global Health Program of the Bill and Melinda Gates Foundation through grant 50728 and the Knowledge for Change Program at the World Bank. The findings, interpretations, and conclusions expressed in this article are those of the authors and do not necessarily represent the views of the World Bank, its executive directors, or the governments they represent. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

† Go to http://dx.doi.org/10.1257/aer.20151138 to visit the article page for additional materials and author disclosure statement(s).

if optimal care requires the potential denial of services that patients value (such as steroids or antibiotics), market-based health care may over-respond to demand, leading to socially inefficient provision (Prendergast 2003). Partly as a result of these considerations, the default policy approach to delivering health care for the poor in most low-income countries is through free or nominally priced medical care in publicly-run facilities staffed by qualified doctors and nurses, who are paid a fixed salary (World Bank 2003).

However, for primary care services a significant fraction of households in low-income countries choose to visit fee-charging health care providers in the private sector; in rural India (the focus of our study), their market share exceeds 70 percent.[1] This is surprising for two reasons. First, private health care providers in India face little de facto regulation and most have no formal medical training (Rohde and Viswanathan 1995; Banerjee, Deaton, and Duflo 2004; CPR 2011). Second, while the high use of the private sector could, in part, reflect the absence of public options, this cannot be the only explanation. In our data from rural India, the private sector share of primary care visits (constructed from a household census) is 83 percent even in markets with a qualified public doctor offering free care through public clinics, and 60 percent of primary care visits in these markets are made to private providers with no formal qualifications.

The high market share of unqualified private health care providers raises a number of questions about the functioning of health care markets in low-income settings. First, why would people choose to pay for care from (mostly) unqualified providers when public clinics are staffed with qualified doctors who offer care at a much lower price? Second, how does the quality of care received vary across public and private health care providers? Third, what does an unregulated health care market reward and how does this compare with the regulated public sector? Specifically, to what extent are prices in the market and wages in the public sector correlated with quality of care? Answers to these questions have been limited by the lack of evidence on the actual quality of care provided in public and private health facilities in low-income settings.[2]

This paper uses data from an audit study conducted in rural areas of the Indian state of Madhya Pradesh (MP) to address this gap. Specifically, standardized (fake) patients (SPs) were coached to accurately present symptoms for three different conditions—unstable angina, asthma, and dysentery in a child (who is at home)—to multiple health care providers. SPs then made over 1,100 unannounced visits to public and private providers of primary health care services and recorded condition-specific metrics of quality of care for each interaction, as well as the price

[1] The market share of private providers is high in many low-income countries: data from the Demographic and Health Surveys (DHS) show that 50 percent of households seeking pediatric outpatient care in Africa and 70–80 percent in India visit the private sector with little variation over the 20 years that these surveys have been collected (IIPS 2007; Grépin 2014). The World Health Surveys include adult morbidity and here the numbers vary from 30 percent in sub-Saharan Africa to between 70 and 80 percent in India (Wagstaff 2013).

[2] Earlier work has highlighted the problem of low doctor effort in the public sector (high absence, low time spent with patients) and low training in the private sector (Banerjee, Deaton, and Duflo 2004; Chaudhury et al. 2006; Das and Hammer 2007). The key evidence gap, however, is the lack of credible estimates of the actual quality of care provided in the public and private sector. For instance, Coarasa, Das, and Gummerson (2016) examine 182 cited studies in two systematic reviews of the medical literature and find only one study that adjusts for differences in patients using an audit methodology (as we do here), and no study that adjusts for differences in providers across public and private practices (which we also do here).

charged.[3] The quality of care metrics include the providers' adherence to a checklist of questions and examinations deemed essential for reaching a correct diagnosis in each case, their likelihood of pronouncing a correct diagnosis, and the appropriateness of the treatments.

We present results from two sets of comparisons. First, we sent SPs to a (nearly) representative sample of public and private health facilities on a walk-in basis, and we use these data to compare the typical patient experience across public and private clinics. However, these differences reflect variation in both provider composition, and differential incentives across public and private clinics. To isolate the effect of practicing in the private sector holding provider characteristics constant, we identified the private practices of qualified public doctors (the majority of whom have one) and sent SPs to present the same medical case to the same set of doctors in both their public and private practices. Our second comparison uses this "dual practice sample" and compares the quality of care across the public and private practices of the same doctors on the same set of cases.

We report three main findings. First, while the majority of private providers in the representative sample have no medical qualifications, they exerted significantly higher effort than public providers and performed no worse on diagnosis and treatment. Private providers spent 1.5 minutes more with patients (62 percent more) and completed 7.4 percentage points more on a checklist of essential history and examination items (47 percent more) than public providers. They were equally likely to pronounce a correct diagnosis (only 4 percent of public providers do so), to offer a correct treatment (27 percent of public providers do so), and to offer clinically unnecessary treatments (provided by 70 percent of public providers). These differences do not reflect high patient loads and waiting times in the public sector; neither do they reflect inadequate equipment and facilities. The results hold even after controlling for these factors and after including market fixed effects.

Second, in the dual practice sample the same doctors spent more time with SPs, completed more items on the checklist, and were also more likely to offer a correct treatment in their private practices, relative to their public practices. Notably, we do not find evidence of differential over-treatment under market incentives, with equivalently high rates of unnecessary treatments, use of antibiotics, and total number of medicines in both types of practices. These differences are *conditional* on seeing the doctor and therefore understate the difference in the quality of patient experiences across public and private practices of the same doctor, because the expected number of trips to the clinic to see a qualified doctor is considerably higher in the public practice due to high doctor absence rates.

Third, we find a positive correlation between the fees charged by private providers and measures of quality such as the time spent, the fraction of checklist items completed, and likelihood of providing a correct treatment. However, we also find a positive correlation between prices and the total number of medications given, including unnecessary treatment. In the public clinics, SPs were provided free or nominally priced care. Since there is no variation in prices, we examine the correlation between

[3] Typically used in medical education, SPs are coached to consistently portray a medical case and all of its physical and psychosocial aspects. When used to evaluate care in hospitals and clinics, they are also trained to accurately recall all aspects of their interactions with the provider. See details in Section II.

doctors' compensation and quality of care and find no correlation between salaries (or desirability of posting) in the public sector and any measure of quality of care delivered.

The main limitation of the SP method is that only a few types of cases can be presented. We therefore complement the SP results with direct observations where enumerators recorded observable details of provider-patient interactions from a full day of within-clinic observations of each provider in our sample (as in Das and Hammer 2007) after SP data collection was completed. We find very similar differences between public and private providers on common measures of quality of care across *all* cases and patient interactions, suggesting that our SP results are likely to be externally valid across a much broader range of cases in this setting.

To help interpret our results, we develop a theoretical framework that models provider-patient interactions in two stages: consultation and treatment. The main insight of the model is that while providers will typically exert more effort in their private practice, the effect on overall patient health is ambiguous. If the default effort level of doctors under low-powered incentives is reasonably high, the marginal gain in diagnostic precision from additional effort in private practice is outweighed by the costs of over-treatment induced by market incentives. On the other hand, if the default effort level is low (as in our setting), the reverse may be true, with better patient health outcomes under market incentives.

Our methodological contribution helps address the fundamental problem of inferring quality in health care, where the optimal action is patient and condition specific, and inefficiencies include undertreatment, overtreatment, or both (Pauly 1980). Specifically, there are four advantages to the use of unannounced SPs relative to existing measures in the literature, which are based on tests of provider knowledge or observation of medical practices.

First, the use of SPs ensures a common set of patient and illness characteristics, which limits concerns about differential patient sorting across clinics on the basis of personal or illness characteristics, as might be the case when observing real patient-provider interactions. Second, the SP method allows us to objectively score the quality of care using condition-specific metrics (checklist completion, diagnosis, and treatment) because we *know* the actual illness being presented and the optimal care associated with the case. In the case of observations with real patients, we would observe only the presenting symptoms and would have to speculate about the true underlying illness.[4] Third, we are able to observe prices charged for completed transactions, which allows us to study the extent to which the unregulated market rewards quality and which improves upon audit studies in other settings that obtain price quotes but do not complete the purchase.[5] Finally, Hawthorne effects are not

[4]Medical vignettes, which measure provider knowledge, also allow for standardization of case-mix and knowledge of the actual illness underlying the presented symptoms, but do not measure actual provider practice, which has been shown to differ markedly from provider knowledge in multiple contexts (Rethans et al. 1991; Leonard and Masatu 2005; Das and Hammer 2007).

[5]For instance, first price offers can be very different from the price of the completed transaction if the distribution of willingness to pay is different across populations. See for instance, Ayres and Siegelman (1995) and Goldberg (1996) for an example of how the lack of completed sales data can lead to misleading conclusions in audit studies of car sales. In our case, the "sale" is always completed as the SP leaves only after the provider has completed the interaction and the price has been paid.

a concern in the SP context because providers do not know that they are being observed.

Substantively, the advances in measurement above combined with our ability to observe the *same* doctor across public and private practices allow us to provide the first direct comparison of the quality of care across public and private sectors.[6] We also provide the first evidence on how market prices for health care behave in an unregulated setting and show that there is a positive correlation between price and checklist completion (and correct treatment), but also between price and unnecessary treatments. This suggests that while unregulated market prices do reflect some information on the quality of care, patients cannot evaluate whether they are being over-treated and charged for unnecessary treatments.

These findings are consistent with the broader empirical literature on credence goods that has demonstrated over-provision of services to the detriment of customer welfare in settings ranging from cesarean sections to car repairs and cab rides for tourists (Gruber and Owings 1996; Schneider 2012; Balafoutas et al. 2013). However, inefficiencies in market provision do not imply that public provision will do better, and a key contribution of our paper is the ability to compare public and private provision of a canonical credence good such as health care.

Combined with the theoretical framework, our results suggest that in settings of poor governance and administrative accountability in the delivery of primary health care services through the public sector (Banerjee, Deaton, and Duflo 2004; Banerjee, Duflo, and Glennerster 2008), market-based provision of health care may present a legitimate alternative in spite of its many theoretical (and empirical) weaknesses. Further, while public health care is free to the consumer, it is not free to the taxpayer. We calculate the per-patient cost in the public sector and conservatively estimate it to be four times higher than the fees charged by private providers in our sample. Thus, the unregulated private market for health care, which is mainly staffed by unqualified providers, appears to deliver higher provider effort and comparable quality of care, at a much lower cost per patient. Our results have direct implications for global policy debates on the organization and delivery of health care services in low-income countries with low state capacity to deliver effective oversight over public health care systems. We discuss these along with caveats in the conclusion.

The rest of this paper is organized as follows. Section I describes health care provision in rural India and Madhya Pradesh; Section II describes the standardized patient (SP) methodology, sampling, data, and measures of health care quality; Section III presents results on quality of care; Section IV covers pricing and cost-effectiveness; Section V discusses robustness to alternative explanations; Section VI presents a theoretical framework to interpret our results; and Section VII concludes with a discussion of policy implications and caveats.

[6]Our approach parallels a literature that tests for moral hazard in developing-country labor markets by comparing worker effort and output under different contractual arrangements (Shaban 1987; Foster and Rosenzweig 1994), and extends it to a credence good setting where output is harder to measure for both customers and researchers, and where there is substantial direct provision of the good by the public sector.

I. Context

A. *Health Care in Rural India*

Health care in India is delivered by both public and private clinics and hospitals. In the public sector, patients can obtain primary care on a walk-in basis in facilities differentiated by their level of specialization ranging from district hospitals and community health centers (CHCs) to public health centers (PHCs) and sub-centers.[7] PHCs, CHCs, and hospitals are supposed to be staffed with trained doctors, who are expected to make diagnoses and either treat or refer patients as appropriate (although in practice, doctor positions are often vacant). Sub-centers are supposed to be staffed with qualified nurses with doctors visiting on a fixed rotation. Most doctors hold a Bachelor of Medicine and Bachelor of Surgery (MBBS) degree, the rough equivalent of an MD in the United States, and receive a fixed salary from the government, with no variable compensation based on either patient load or quality of care.[8]

Consultations in public clinics are provided on a walk-in basis during opening hours (appointments are rarely used), and are free or nominally priced. Patients are also supposed to receive free medication, if available. Although a federally-funded insurance program for inpatient hospital care was introduced in 2007, the tax-funded public system of care was the only source of (implicit) public insurance for primary care.

In theory, public facilities are accountable to administrative norms and procedures (documented in the Civil Service Codes for each state). In practice, administrative accountability of public health care providers is weak. Nationwide, doctor absences in public clinics averaged 43 percent on any given day in 2003 and 40 percent in 2010 (Muralidharan et al. 2011; CPR 2011). These absences do not occur on predictable days or hours (Banerjee, Deaton, and Duflo 2004) and they are not easy to address at a system-level (Banerjee, Duflo, and Glennerster 2008; Dhaliwal and Hanna 2017). When asked about adherence to administrative rules, more than 80 percent of public sector doctors agree that the rules and norms are frequently flouted and that appropriate "payments" can allow providers to circumvent disciplinary proceedings, even for grave negligence (La Forgia and Nagpal 2012).

While official policy documents of the Government mainly focus on improving the public system of primary health care (Planning Commission of India 2013), data from household surveys consistently show that the fee-charging private sector accounts for over 70 percent of primary care visits (IIPS 2007; Selvaraj and Karan 2009; CPR 2011). Barriers to entry for private health care providers are low. Provider qualifications range from MBBS degrees to no medical training at all, and clinics can range from well-equipped structures to small one-room shops, the provider's residence, or the patients' home for providers that make home visits.

[7] Official guidelines stipulate that there should be a sub-center for every 5,000 people, a primary health care center for every 25,000 people, and a community health center for every 100,000 people.

[8] India also recognizes medical degrees from alternative schools of medicine including Ayurveda, homeopathy, and Unani. However, providers with these qualifications are only licensed to prescribe medication in line with their training and are not licensed to prescribe allopathic medicine. They also are not typically posted in the frontline health care system of PHCs, CHCs, and district hospitals that prescribe allopathic medicine.

Providers operate on a fee-for-service basis, and prices often include the cost of medicines. While providers operating without a medical license are not legal and face the threat of being shut down, they have come to be the dominant source of care in these markets (as the data below will show).

B. *Sampling of Health Care Markets and Summary Statistics*

We carried out the SP study in the Indian state of Madhya Pradesh (MP), one of India's poorer states, with a GDP/capita of ~ \$600/year (or ~ \$1,500/year in PPP terms) in 2010–2011 (the period of the study). We first drew a representative sample of 100 villages across 5 districts, stratified by geographic regions and an index of health outcomes. We then conducted a household *census* in these villages, where respondents named all providers from whom they sought primary care in the previous 30 days and their locations (including providers practicing outside the village). We then surveyed all providers in all of these locations, regardless of whether or not the providers themselves had been mentioned in the sample villages, thereby obtaining a census of all providers in the health care market that catered to sampled villages (see online Appendix Figure A.1).

Table 1 (columns 1–3) presents summary statistics based on the provider census (panel A) and the household census (panel B) in these markets; columns 4–6 compare villages sampled for the SP study to the representative villages. The table highlights three key features of health markets in rural India. First, villages are served by a large number of providers once the health market is correctly accounted for by including locations that are nearby but outside village boundaries. There are 11 primary care providers per market and 46 percent of households reported visiting a primary care provider in the 30 days prior to the survey.

Second, the majority of providers are private (7 out of 11 or 64 percent), and they account for 89 percent of household visits; excluding paramedical public health workers (typically responsible for preventive, maternity, and child care) increases the fraction further to 93 percent. The share of visits to private providers (with or without qualifications) is 88 percent when there is a public provider in the market, and is 83 percent even when there is a public MBBS doctor in the same market.

Third, 46 percent of all providers and 70 percent of all private providers (5.4 per village) have no formal medical training, yet they account for 77 percent of household visits. There is less than one MBBS doctor per market, and one is rarely available within the village. The distribution of MBBS providers is uneven. Only 30 percent of all villages have recourse to an MBBS provider (public or private) in their market, and only 5 percent have one within village boundaries. Private unqualified providers remain the dominant providers of care in most settings, accounting for 74 percent of all visits even when there is a public provider in the same market, and 60 percent even when there is a public MBBS doctor in the same market.[9] MBBS doctors account for only 4 percent of all patient interactions (panel B).

[9] Note that even public facilities have many unqualified providers. While these are typically support staff (who are only supposed to assist a qualified doctor), we find that it is very common for these staff to act as the main health care providers in public clinics and prescribe medication (given high doctor absence rates).

Table 1—Health Market Attributes

	Madhya Pradesh (5 districts, 100 markets)			SP sample villages (3 districts, 46 markets)		
	All (1)	Inside village (2)	Outside village (3)	All (4)	Inside village (5)	Outside village (6)
Panel A. Composition of markets based on census of providers						
Total	11.68	3.97	7.71	16.02	4.65	11.37
	(12.06)	(4.49)	(12.17)	(15.81)	(5.41)	(16.42)
Public MBBS	0.45	0.05	0.40	0.50	0.02	0.48
	(0.97)	(0.22)	(0.93)	(1.11)	(0.15)	(1.11)
Public alternative qualification	0.22	0.07	0.15	0.24	0.07	0.17
	(0.48)	(0.29)	(0.39)	(0.52)	(0.33)	(0.44)
Public paramedical	1.58	1.13	0.45	1.98	1.30	0.67
	(1.90)	(1.46)	(1.33)	(2.12)	(1.49)	(1.59)
Public unqualified	1.71	0.68	1.03	2.07	0.67	1.39
	(1.75)	(1.04)	(1.54)	(2.05)	(1.12)	(1.94)
Total public	3.96	1.93	2.03	4.78	2.07	2.72
	(3.20)	(2.28)	(2.63)	(3.53)	(2.45)	(3.17)
Private MBBS	0.40	0.00	0.40	0.59	0.00	0.59
	(1.57)	(0.00)	(1.57)	(2.15)	(0.00)	(2.15)
Private alternative qualification	1.92	0.23	1.69	2.67	0.33	2.35
	(3.65)	(0.66)	(3.65)	(4.86)	(0.90)	(4.89)
Private unqualified	5.40	1.81	3.59	7.98	2.26	5.72
	(6.01)	(2.23)	(6.14)	(7.88)	(2.74)	(8.32)
Total private	7.72	2.04	5.68	11.24	2.59	8.65
	(10.54)	(2.69)	(10.81)	(14.31)	(3.38)	(14.87)
Panel B: Composition of demand from census of households in sampled villages						
Fraction of households that visited a provider in last 30 days	0.46			0.58		
	(0.50)			(0.49)		
Fraction provider visits inside/outside village		0.66	0.34		0.69	0.31
		(0.47)	(0.47)		(0.46)	(0.46)
Distance traveled to visited provider (km)	1.61	0.40	3.83	1.37	0.38	3.51
	(2.14)	(0.65)	(2.14)	(2.37)	(1.16)	(2.84)
Fraction of visits to MBBS doctor	0.04	0.01	0.09	0.02	0.00	0.06
	(0.19)	(0.09)	(0.29)	(0.13)	(0.00)	(0.23)
Fraction of visits to private sector	0.89	0.92	0.85	0.96	0.97	0.93
	(0.31)	(0.28)	(0.36)	(0.21)	(0.18)	(0.26)
Conditional on public availability	0.88	0.89	0.83	0.95	0.96	0.91
	(0.33)	(0.31)	(0.38)	(0.22)	(0.20)	(0.28)
Conditional on public MBBS availability	0.83	0.84	0.79	0.93	0.98	0.90
	(0.37)	(0.36)	(0.41)	(0.25)	(0.15)	(0.30)
Fraction of visits to unqualified providers	0.77	0.87	0.55	0.82	0.89	0.64
	(0.42)	(0.34)	(0.50)	(0.39)	(0.31)	(0.48)
Conditional on public availability	0.74	0.82	0.54	0.81	0.86	0.64
	(0.44)	(0.38)	(0.50)	(0.39)	(0.35)	(0.48)
Conditional on public MBBS availability	0.60	0.77	0.38	0.66	0.81	0.39
	(0.49)	(0.42)	(0.48)	(0.47)	(0.39)	(0.49)
Panel C. Sample characteristics from household census of provider choice						
Number of villages	100			46		
Average village population	1,149			1,199		
Average number of households per village	233			239		
Number of reported provider visits	19,331			12,122		
Average number of visits per household per month	0.83			1.10		

Notes: Standard deviations in parentheses. The number of providers available to a village was determined by a provider census, which surveyed all providers in all locations mentioned by households in 100 sample villages, when asked where they seek care for primary care services, regardless of whether or not the particular provider was mentioned by households. Unqualified providers report no medical training. All others have training that ranges from a correspondence course to a medical degree. Outside villages are typically adjacent villages or villages connected by a major road. The 30-day visit rate was calculated from visits to providers reported by households in a complete census of households in the 100 sample villages. The type of provider they visited was determined by matching reported providers to providers surveyed in the provider census.

Source: Authors' calculations

II. Measuring Health Care Quality Using Standardized Patients

A. *The Standardized Patient (SP) Methodology*

Used routinely in the training and evaluation of medical students in high-income countries, including the United States, SPs are highly trained "fake patients" who present symptoms of an illness to a physician like any other normal patient. Details of the interactions when SPs are unknown or unannounced to the providers beforehand can be used to evaluate the quality of care received by a typical patient (Rethans et al. 1991). SPs are coached to present their initial symptoms and answer any questions that the physician may ask as part of history taking, in a manner consistent with the underlying condition. We followed the same method (adapted to local conditions) and sent unannounced SPs to health care providers in our sample during the course of a normal working day.

A total of 15 SPs were recruited from the districts where the study was conducted. Using a team that included a professional SP trainer, two medical doctors, and a medical anthropologist familiar with local forms of presenting symptoms and illnesses, SPs were coached to accurately and consistently present one of three cases: unstable angina in a 45-year-old male, asthma in a 25-year-old female or male, and dysentery in a child who was at home presented by the father of the child (see Das et al. 2012 and online Appendix B for details on SP protocols).[10] SPs visited sampled providers, who did not know they were receiving standardized patients and therefore should have treated them as new patients.[11] After the interaction, SPs were debriefed within an hour with a structured questionnaire that documented the questions and examinations that the provider completed or recommended, the treatments provided, and any diagnoses offered. The SPs retained any medicines dispensed in the clinic and paid all fees charged by providers at the end of the interaction.

The SPs depicted uncomplicated textbook presentations of the cases, and a panel of doctors who advised the project concurred that appropriate history taking and examinations should lead providers toward the correct diagnosis and treatment. Cases were specifically chosen so that the opening statement by the SPs would be consistent with multiple underlying illnesses, but further questioning should have led to an unambiguous (correct) diagnosis. This allows us to measure provider quality through adherence to an essential checklist of questions and examinations that would allow them to accurately make a diagnosis and provide a correct treatment. We also chose these cases since they represented conditions with high or growing incidence in India and other middle- and low-income countries, and they minimized risk to SPs that could arise from unsafe invasive examinations, such as a blood test with an unsterilized needle.

[10] Das et al. (2012) discusses the SP methodology in further detail and presents summary statistics on overall quality of care in this setting. The current paper focuses on the economics of unregulated health care markets and we do not replicate the analysis in Das et al. (2012). See online Appendix B for further details on how the SP method was implemented, including further discussion on the choice of cases and their relevance. Details on case presentations and instruments are posted on www.healthandeducationinindia.org.

[11] The research ethics board of Innovations for Poverty Action approved this design following a successful pilot in Delhi, where the detection rate of SPs was extremely low even among a set of doctors who were informed that they would receive an SP at some point in the next month.

In these cases the role of suitable medical advice was important because real patients would be unlikely to be able to categorize the symptoms as "life threatening" or "potentially non-harmful" and triage themselves into clinics or hospitals. For instance, the SP with unstable angina complains of chest pain which, even in countries with advanced health systems, is often mistaken by patients as arising from heartburn, exertion, or muscle strain.[12] Similarly, wheezing and shortness of breath in asthma may arise from short-term allergies to environmental contaminants. Finally, for any child with diarrhea, a key contribution of a health care provider is to assess whether the symptoms reflect a bacterial or viral infection (and thus whether the patient requires antibiotics) and the degree of dehydration—each of which may be difficult for parents to assess.

B. *Health Care Provider Sampling and Summary Statistics*

Our study first uses the census of health care providers described earlier to construct a near representative sample of public and private health care providers in three of the five sampled districts in rural MP. While our SPs were recruited from the districts in our sample, they were never residents of the villages where they presented themselves to health providers. Since providers in rural areas might know their patients, the SPs had to justify their presence in the area by mentioning, for example, work-related travel or visits to relatives. For such excuses to be plausible, our final sample dropped villages that could not be accessed by paved roads and comprised a total of 46 villages across three districts. While these sampled villages have more providers on average than the entire representative set of villages, there is no difference in the composition of providers across the frame and sample (Table 1).

Since SPs visited clinics to obtain primary care, we excluded community health workers, midwives, and providers that only made home visits. We then sampled all public clinics (some large ones were sampled twice), and a maximum of six private providers in each market for a total of 235 clinics, and SPs completed interactions with 224 providers.[13]

Data from this "representative sample" allow us to compare care provided across typical public and private clinics in rural MP (all estimates are re-weighted by the inverse of the sampling probabilities to provide population representative averages). However, this comparison would reflect a combination of any compositional differences among providers across public and private clinics, as well as the effect of practicing in the private sector.

To isolate the role of private sector practice, we identified the universe of public MBBS doctors posted to PHCs and CHCs from all five study districts, even if these clinics were not located in the village-based sampling scheme. We then identified the private practices of these doctors (we found a private practice for 61 percent). We sampled and successfully administered SP visits to 116 public MBBS doctors. Our "dual sample" consists of the 91 doctors in this MBBS sample who also have

[12] The REACT study in the United States found that many chest pain patients delayed calling 911 because they confused their symptoms with heartburn (Faxon and Lenfant 2001).

[13] In one case, a sampled village was near a market with over 100 different health care providers. In this one case, we sampled over 20 private providers. See online Appendix A for further details on sampling.

a private practice, and for 70 of these, SPs presented cases in both their public and private practices. The "dual sample" enables a comparison of the quality of care provided by the same doctor on the same case across his public and private practices. SP completion rates in the dual sample were higher in the private (92 percent) compared to public practices (78 percent), due to higher doctor absence rates in their public practice, leading to non-completion despite multiple attempts. We show that all our results are robust to adjustments for differential non-completion rates (see Section IIIF and online Appendix D.1).

Note that in the representative sample, the unit of analysis is the *clinic* and the SP experience is recorded based on whoever they saw in the clinic. In the dual sample, the unit of analysis is the *doctor* and the SP made repeat visits to see the sampled doctor if needed (especially in the public practice). Online Appendix A and Tables A.1 and A.2 provide further details on the sampling and construction of the representative and dual samples.

Table 2 (columns 1–3) provides summary statistics for the representative sample of providers. The providers are mostly middle-aged men and just under 60 percent have completed 12 or more years of education (Table 2, panel A). Their practices have been open for 13–15 years, and private and public providers self-report an average of 16 and 28 patients per day, respectively. Most practices (82 percent of private and 100 percent of public) dispense medicines in the clinic itself and are equipped with the infrastructure and medical devices required for routine examinations, such as stethoscopes and blood pressure cuffs. In the representative sample, public providers are more likely to have an MBBS degrees (26 percent versus 8 percent). Private providers charged an average of Rs 51 per interaction. Consistent with nominally priced public care, our SPs paid Rs 3.7 on average in public clinics.

Strikingly, 70 percent of private providers in the representative sample report no formal or unverifiable medical qualifications. However, most of them do have noncredentialed medical training. Online Appendix Table A.3 presents details of medical training in the representative sample, and we see that 86 percent of unqualified providers report having received additional training, with the average duration being 32 months. Similarly, 75 percent of providers with unrecognized qualifications report noncredentialed medical training averaging 37 months. The most common form of training is from being an assistant in another doctor's practice. Field interviews suggest that these providers also receive informal continuing medical education from pharmaceutical sale representatives. Thus, while they have no formal qualifications and are not legally licensed to practice, these providers do have considerably greater medical knowledge than a lay person and command considerable credibility in their communities (as suggested by their high market share).

Column 4 presents summary statistics on the universe of public MBBS doctors, while columns 5–7 present these for the 88 public MBBS doctors in the dual sample and test if they are comparable. Overall, doctors with and without dual practices are similar on observable characteristics, but the former have a longer tenure at their current location. There is no significant difference in the equipment reported across these practices (Columns 8–10), although the overall number of patients seen is higher in the public practice and the fees charged are higher in the private practice.

Table 2—Characteristics of Providers and Practices Where SPs Were Administered

	Representative sample (3 districts)			Representative sample of public MBBS providers (5 districts)				Dual practice sample (5 districts)		
	Public (1)	Private (2)	p-value of (1)–(2) (3)	All public (4)	Non-dual public (5)	Dual public (6)	p-value of (5)–(6) (7)	Public (8)	Private (9)	p-value of (8)–(9) (10)
Panel A. Provider characteristics										
Age of provider	46.92	43.51	0.10	44.52	44.74	44.43	0.89			
Is male	0.86	0.96	0.02	0.87	0.96	0.84	0.10			
More than 12 years of basic education	0.58	0.52	0.48	0.64	0.52	0.69	0.09			
Has MBBS degree	0.25	0.07	0.00	1.00	1.00	1.00				
Has alternative medical degree	0.11	0.21	0.18	0.00	0.00	0.00				
Has no formal medical qualification	0.61	0.68	0.42	0.00	0.00	0.00				
Has noncredentialed medical training	0.63	0.78	0.05	0.23	0.22	0.23	0.96			
Number of practices	1.14	1.07	0.21	1.83	1.16	2.13	0.00			
Tenure in years at current location	15.22	13.70	0.42	6.15	5.11	6.56	0.28			
Panel B. Clinic characteristics										
Dispense medicine	1.00	0.81	0.00							
Consultation fee (Rs)	3.65	51.24	0.00	3.75	3.15	3.92	0.00	3.92	57.93	0.00
Number of patients per day (self reported in census)	28.06	15.74	0.00	31.85	31.30	35.00	0.74	35.00	17.59	0.07
Number of patients per day (from physician observations)	5.72	5.75	0.98	16.04	13.72	16.86	0.31	16.86	5.63	0.00
Electricity	0.94	0.95	0.93	1.00	1.00	1.00		1.00	1.00	
Stethoscope	0.97	0.94	0.47	1.00	1.00	1.00		1.00	1.00	
Blood pressure cuff	0.83	0.75	0.34	1.00	1.00	1.00		1.00	1.00	
Thermometer	0.94	0.92	0.64	0.97	0.94	0.98	0.20	0.98	0.97	0.63
Weighing scale	0.86	0.52	0.00	0.94	0.94	0.94	0.96	0.94	0.82	0.04
Hand-wash facility	0.89	0.81	0.30	0.84	0.84	0.85	0.93	0.85	0.81	0.56
Number of providers	36	188		103	31	72		72	84	

Notes: Standard deviations are in parentheses. Unit of observation is a provider. The dual practice sample consists of providers who received a standardized patient in both their public and private practices. Provider mapping and complete provider census yielded information about whether or not a provider operates more than one practice. The representative sample did not employ the intense reconnaissance to find both the public and private practices of the same provider, and thus the proportion of dual practice providers can be considered self-reported. In the dual practice sample, however, the existence of additional medical practices was verified by repeated observation. Alternative qualifications are as follows: BAMS, BIMS, BUMS, BHMS/DHMS, DHB, BEHMS, BEMS, B.Sc. Nursing/M.Sc. Nursing, B.Pharma/M.Pharma. In the public sector of the representative sample, there are 3 providers with BAMS and 1 with B.Pharma/M.Pharma. In the private sector, there are 21 with BAMS, 9 with BHMS/DHMS, 3 each with BIMS and DHB, 2 with B.Pharma/M.Pharma, and 1 with BUMS. No medical training includes providers with unverifiable degrees and providers who self-reported no formal training. In the public sector of the representative sample, there are 22 with no formal qualifications and 5 who reported other degree. In the private sector, there are 128 with no formal qualifications and 56 who reported other unverifiable degrees. Means for consultation fee were calculated from direct observations of clinical interactions. All other variables derive from a survey administered during the census of providers.

Source: Authors' calculations

We randomly assigned three SPs to each sampled clinic in the representative sample, one presenting each of the three cases. For the dual sample, we sent SPs presenting the asthma and dysentery cases to both practices of the same provider.[14] Since the rarity of

[14] Since we had 15 SPs and 3 cases, we made sure that the same case was presented by different SPs in the public and private practices. To ensure that our standardized patients saw the sampled provider when (s)he visited the public clinic and not a substitute, we first interviewed all providers in their private practices or residences without revealing that we knew they also worked in the public sector, and we obtained either their photograph or a detailed description of their physical appearance. SPs portrayed a dummy case (e.g., headache) if the doctor was absent when they visited the public clinic, and we sent in other SPs on our subsequent attempts. As we discuss later, it took

unstable angina could have raised suspicions if providers saw two travelers presenting the same case (even though visits were typically separated by a few weeks), we randomized the providers into two groups: one that received an unstable angina patient in his/her private practice and another that received the case in the public clinic. We show that the randomization was valid in online Appendix Table A.4.

C. *Measuring Quality of Care*

We use three measures of quality of care. Our first metric is the extent to which the provider adhered to a checklist of questions and examinations required for making a differential diagnosis on each of the presented cases. For instance, these questions and exams would allow a doctor to distinguish between heartburn (that has gastrointestinal origins) and a heart attack, or between viral diarrhea and dysentery. These items represent a parsimonious subset of the Indian government's own guidelines, and the list we use was developed by a panel of Indian and American doctors (the items are described for each case in online Appendix Table A.5).[15] While the most transparent measure of checklist adherence is the percentage of checklist items completed, we also compute an index score using item response theory (IRT), which gives more weight to items that discriminate better among providers. Developed in the context of educational testing, IRT allows us to create a composite measure of provider quality based on questions asked across all three cases, with lower weights on checklist items that are less essential and higher weights on more essential questions that do a better job of discriminating between low and high quality providers (see Das and Hammer 2005 for details). We report both measures in our analysis.

Second, we examine diagnoses: whether one was provided and whether it was correct. We only classify a diagnosis as correct if the provider specified the actual ailment that the SP presented or a functional equivalent. Online Appendix Table A.5, panel B presents the diagnoses that were considered correct for each case, and also provides a sense of the wide range of incorrect diagnoses that were seen in practice.

Third, we evaluate the quality of treatment provided. SPs noted all treatment instructions received and retained all prescriptions and medication dispensed in the clinic. These were then classified as correct, palliative, or unnecessary/harmful, based on inputs from our panel of doctors, pharmacists, and a pharmaceutical company (see online Appendix B.4 for details; online Appendix Table A.5, panel C lists specific treatments in each category). Since providers can dispense or prescribe multiple medicines, we classify each medicine as correct, palliative, or unnecessary/harmful and thus allow the total treatment protocol to be classified into multiple categories at the same time.

Correct treatment refers to a treatment that is clinically indicated for the specific case and that would relieve/mitigate the underlying condition. Palliative treatments

significantly more trips to complete an SP case in the public practice relative to the private one, due to the high rates of provider absence in the public practice.

[15] The Indian government's National Rural Health Mission (NRHM) has developed triage, management, and treatment protocols for unstable angina, asthma, and dysentery in public clinics, suggesting clear guidelines for patients presenting with any of these conditions. The checklist we use is more parsimonious. If we had used the more extensive checklist and asked the SPs to recall adherence to more items, it is likely that checklist adherence would be lower than the numbers that we document.

are those that may provide symptomatic relief, or treatments where the providers correctly identified which system was being affected, but which on their own would not cure the patient of the condition that was being presented: for example, allergy medicine for the asthma patient. Treatments classified as unnecessary/harmful were neither correct nor palliative. We group these two potentially distinct categories together because it was difficult to achieve consensus among doctors on what should be considered harmful. Some, for example, would consider antibiotics for the unstable angina patient unnecessary. Others took a longer view with antibiotic resistance in mind and considered it as ultimately harmful. However, none of the treatments we observed were directly contra-indicated, and hence most of these represent unnecessary treatments as opposed to directly harmful ones.[16]

However, even after classifying all medicines as correct, palliative, and unnecessary/harmful, there are two challenges in coding the "correctness" of a treatment. The first is: How should we interpret a referral when incentives are very different? In some cases, this may be a good thing (if, for example, the provider refers a heart attack patient to a hospital). In other cases, a "referral" may simply reflect a provider who deflected the case without directing the patient usefully.[17] Since we did not send the SPs to the place that was referred, there is no obvious way of coding the quality of referrals. We therefore try to be conservative in our main analysis and do not treat referrals as correct treatments. When we repeat the analysis treating referrals as correct in the angina case, our results are unchanged (see below).

A second challenge arises from the proxy nature of the dysentery case. Many providers did not provide a treatment because the child was not presented and instead asked to see the child. We therefore report results for "checklist completion" using all three cases, but drop the dysentery case for "diagnosis" and "treatment" because the patient (the sick child) was not actually presented for this case. All results are robust to dropping the case completely.

III. Results: Quality of Care across Public and Private Providers

A. *Estimation Framework*

Our main interest is in estimating differences in the quality of care that patients received from providers in the public and private sectors. In the representative sample, we estimate

$$q_{(i(scp)m)} = \beta_0 + \beta_1 \text{Private}_{ip} + \beta_2 X_p + \delta_s + \delta_c + \delta_m + \epsilon_{i(scp)m}, \tag{1}$$

where we regress each measure of quality q (checklist completion, diagnosis, and treatment) in interaction i between a standardized patient s presenting case c and

[16] If the overall quality of care were higher, we could have designed the SP case with a patient who is allergic to certain kinds of antibiotics or who is on regular medication for another illness. In this case, many treatments would have been harmful and the case would have required the doctor to watch out for drug interactions. Given the low-level of overall quality of care, designing such an SP case would not have been very useful at discriminating quality because SPs were never asked about existing allergies or whether they were currently taking any medication.

[17] Field notes suggest that this often happened in public clinics where the doctor was absent. The available provider did not ask questions or conduct any examinations, and told the SP to go elsewhere. By necessity, this is coded as a "referral" in our data, although the patient received no information from the interaction.

a provider p in market m on an indicator for the sector (Private), with β_1 being the coefficient of interest. Since we pool cases and SPs and there may be systematic differences across them, all our specifications include SP and case fixed effects (δ_s and δ_c). We report three sets of estimates for each quality measure. First, we include only SP and case fixed effects; then we add market fixed effects so that comparisons reflect relative performance in the same market (note that not all markets had both types of providers); finally, we add controls for provider and practice characteristics X_p, to adjust for observable differences across providers including demographics, reported qualifications, and number of patients waiting during the visit.

While β_1 provides a useful estimate of the differences in quality across public and private providers in a representative sample of providers, it is a composite estimate that includes differences in unobservable provider characteristics, as well as the effect of practicing in the private sector. To isolate the impact of private sector practice, we re-estimate equation (1) in the dual sample that only includes data from the cases where we sent the SPs to the public and private practices of the same MBBS doctor. We report three sets of estimates here as well. First, we include only SP and case fixed effects;[18] then we add district fixed effects (since the dual practice sample was drawn from the universe of public MBBS doctors practicing in each district rather than the universe of providers practicing in sampled villages, as was the case for the representative sample); finally, we include controls for observable differences across the public and private practices of the doctors.

B. *Completion of Essential Checklist of History Taking and Examinations*

Columns 1–3 in Table 3 present results from estimating equation (1) in the representative sample. Our outcome variable is "provider effort," measured by consultation length and checklist completion. While the results are similar across the three specifications, we focus our discussion on the estimates in panel B, because they compare relative performance within the same market (without controlling for provider characteristics), which is the relevant choice set for patients. The base level of effort among representative public providers was low. The average public provider spent 2.4 minutes with the SP in a typical interaction and completed 16 percent of checklist items. Private providers spent 1.5 minutes more per patient and completed 7.4 percentage points more items on the checklist (62 percent and 47 percent more than the public providers respectively). When evaluated on the IRT-scaled score, private providers scored 0.67 standard deviations higher. Figure A.2 shows that time spent with the patient is strongly correlated with the number of checklist items completed, which points to the credibility of the SP presenting the case, as more time spent with the patient led to greater checklist completion.

Columns 4–6 repeat the analysis in the dual sample, with similar results. Public MBBS doctors appear to be more productive than the typical public provider in

[18] Note that we do not include provider fixed effects since the angina case was not presented in both the public and private practices of the same doctor and will drop out if we do so. Since the case was randomly allocated across the public and private practices of the doctor and assignment was balanced on measures of quality of other cases (see online Appendix Table A.4), our estimates will be an unbiased estimate of the average quality difference across the public and private practices of public MBBS doctors. We also estimate equation (1) with provider fixed effects and the results are unchanged.

Table 3—Effort in the Public and Private Sectors

	Representative sample			Dual practice sample		
	Time spent (mins) (1)	Percentage of checklist items (2)	IRT score (3)	Time spent (mins) (4)	Percentage of checklist items (5)	IRT score (6)
Panel A. SP and case fixed effects						
Is a private provider	1.222	6.758	0.551	1.507	8.977	0.755
	(0.250)	(2.488)	(0.212)	(0.298)	(1.935)	(0.207)
R^2	0.305	0.160		0.241	0.220	
Observations	662	662	233	331	331	138
Mean of public	2.388	15.287		1.561	17.720	
Mean of private	3.703	22.302		2.983	28.308	
Mean of sample	3.603	21.764		2.274	23.030	
Panel B. SP, case and market/district fixed effects						
Is a private provider	1.486	7.352	0.668	1.514	8.977	0.759
	(0.333)	(2.705)	(0.277)	(0.298)	(1.922)	(0.207)
R^2	0.391	0.259		0.262	0.234	
Observations	662	662	233	331	331	138
Panel C. SP, case and market/district fixed effects						
Is a private provider	1.246	5.999	0.611	1.485	9.504	0.829
	(0.424)	(2.891)	(0.327)	(0.316)	(2.062)	(0.205)
Has MBBS	−0.156	3.285	0.043			
	(0.638)	(2.589)	(0.257)			
Has some qualification	−0.131	2.518	0.157			
	(0.443)	(1.813)	(0.151)			
Age of provider	−0.004	−0.046	0.000	0.004	−0.066	0.004
	(0.013)	(0.059)	(0.008)	(0.019)	(0.089)	(0.101)
Gender of provider (1=male)	0.653	−0.949	0.212	−0.070	−1.343	−0.288
	(0.428)	(4.207)	(0.327)	(0.437)	(3.306)	(0.309)
Patient load during visit	−0.096	−0.144	0.082	−0.097	−0.225	0.013
	(0.061)	(0.481)	(0.040)	(0.041)	(0.457)	(0.517)
R^2	0.399	0.259		0.278	0.234	
Observations	638	638	221	302	302	126

Notes: For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level, except in IRT score where each observation is a composite provider level score across all cases. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Source: Authors' calculations

the representative sample (many of whom are unqualified) because they complete a slightly higher fraction of checklist items (18 percent) in 35 percent less time (0.8 minutes less). However, this additional productivity is not used to complete more checklist items in the public practice, but rather to reduce the time spent with patients (1.56 minutes versus 2.4 minutes in the representative sample). In their private practices, the same doctors doubled consultation length, completed 60 percent more checklist items, and scored 0.76 standard deviations higher on the IRT-scaled measure of quality. It is worth comparing these differences with those obtained

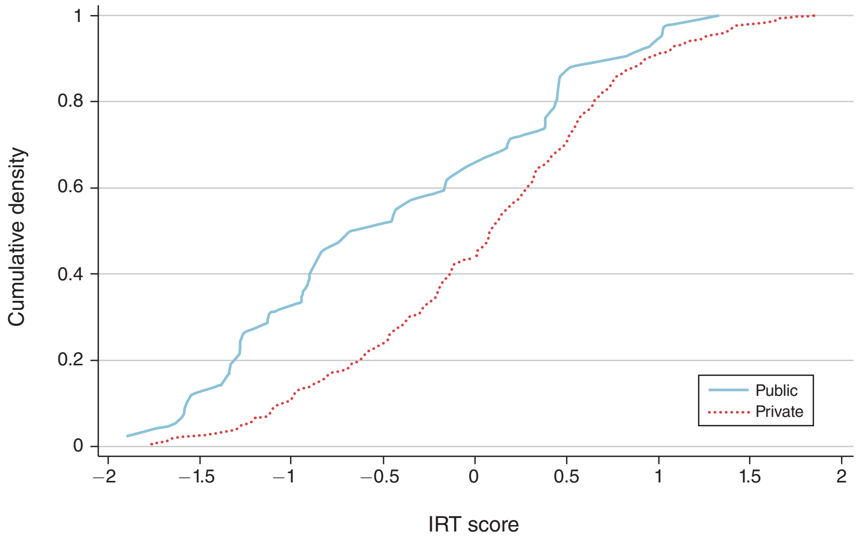

FIGURE 1. CHECKLIST COMPLETION BY PROVIDER TYPE IN THE REPRESENTATIVE SAMPLE

in interventions that are regarded as highly successful. For instance, Gertler and Vermeersch (2013) look at checklist completion as a result of the introduction of performance pay in Rwanda. They find that performance pay increased checklist completion by 0.13 standard deviations; we find that the difference in checklist completion across public and private practices of the same doctor is over five times larger.

These differences are seen clearly in Figures 1–3. Figure 1 plots the cumulative distribution functions (CDF) of the IRT score (based on checklist completion) of public and private providers in the representative sample, Figure 2 does so for the dual sample, and Figure 3 pools all four samples together (Figures A.3–A.5 plot the corresponding distributions). The distribution of checklist completion for private providers first-order stochastically dominates that of the public providers (Figure 1) and the corresponding distribution for the private practices of public providers also first-order stochastically dominates that of their public practices (Figure 2). Finally checklist completion is higher for public MBBS doctors than a representative public provider (as would be expected given that the former are more qualified), but it is lower for the public MBBS doctors even relative to a representative sample of private providers (most of whom are unqualified, Figure 3).

Focusing on individual checklist items (online Appendix Table A.6) shows that private providers in both samples are significantly more likely to perform several items on the checklist on all three cases and are no less likely to perform any of the items (except for one in asthma). In addition to β_1, Table 3 (columns 1–3) also shows that there is no statistically significant correlation between the possession of any formal medical qualification and checklist completion, suggesting that formal qualifications may be a poor predictor of provider effort.

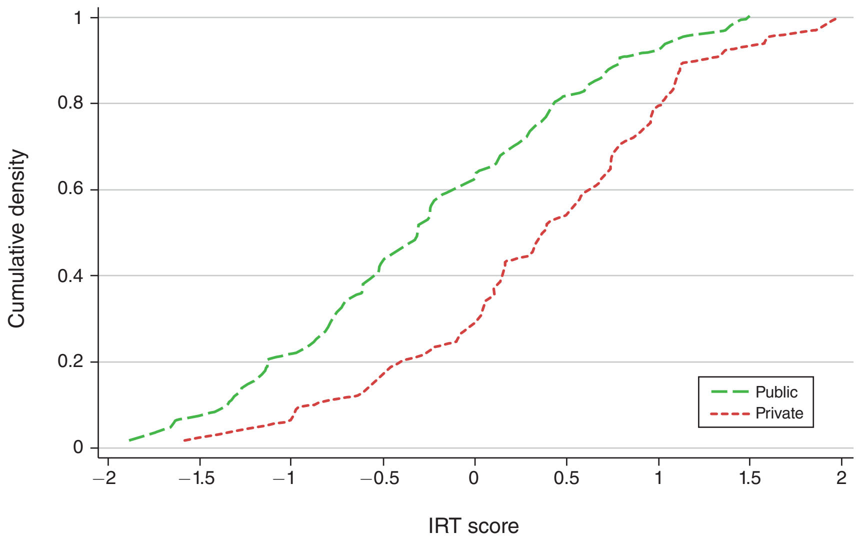

Figure 2. Checklist Completion by Provider Type in the Dual Practice Sample

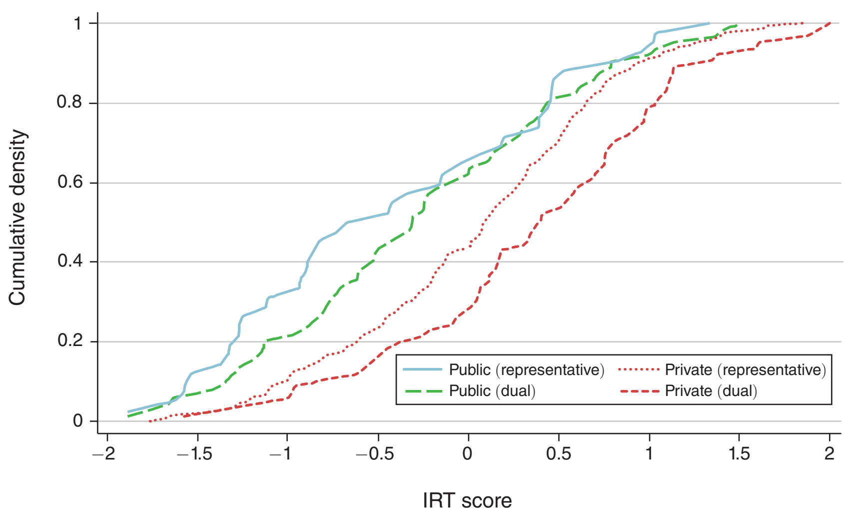

Figure 3. Checklist Completion by Provider Type

C. *Diagnosis*

Results for diagnosis (Table 4) follow the same format as Table 3 but the dependent variables of interest are whether any diagnosis was given and whether a correct diagnosis was given (both conditional and unconditional on uttering a diagnosis). In the representative sample, 26 percent of public providers offer a diagnosis, of whom

TABLE 4—DIAGNOSIS IN THE PUBLIC AND PRIVATE SECTORS (*Unstable Angina and Asthma Cases Only*)

	Representative sample			Dual practice sample		
	Gave diagnosis (1)	Correct diagnosis (conditional) (2)	Correct diagnosis (unconditional) (3)	Gave diagnosis (4)	Correct diagnosis (conditional) (5)	Correct diagnosis (unconditional) (6)
Panel A. SP and case fixed effects						
Is a private provider	0.168	−0.014	0.016	0.095	−0.041	0.023
	(0.052)	(0.057)	(0.022)	(0.066)	(0.102)	(0.049)
R^2	0.130	0.121	0.075	0.130	0.113	0.055
Observations	440	178	440	201	88	201
Mean of public	0.263	0.150	0.039	0.382	0.385	0.147
Mean of private	0.431	0.135	0.058	0.495	0.388	0.192
Mean of sample	0.418	0.135	0.057	0.438	0.386	0.169
Panel B. SP, case and market/district fixed effects						
Is a private provider	0.188	−0.019	0.023	0.092	−0.056	0.025
	(0.061)	(0.071)	(0.027)	(0.068)	(0.107)	(0.049)
R^2	0.218	0.301	0.145	0.150	0.175	0.067
Observations	440	178	440	201	88	201
Panel C. SP, case and market/district fixed effects						
Is a private provider	0.149	−0.046	0.031	0.084	0.017	0.044
	(0.067)	(0.095)	(0.032)	(0.071)	(0.120)	(0.055)
Has MBBS	−0.092	0.108	0.008			
	(0.125)	(0.131)	(0.030)			
Has some qualification	0.023	−0.010	−0.012			
	(0.058)	(0.073)	(0.025)			
Age of provider	−0.002	−0.005	−0.002	0.002	−0.001	0.000
	(0.002)	(0.003)	(0.001)	(0.004)	(0.009)	(0.004)
Gender of provider (1=male)	−0.089	0.272	0.079	−0.125	−0.052	−0.086
	(0.134)	(0.092)	(0.030)	(0.095)	(0.174)	(0.071)
Patient load during visit	−0.003	−0.017	−0.005	−0.017	−0.003	−0.005
	(0.008)	(0.009)	(0.005)	(0.020)	(0.033)	(0.011)
R^2	0.222	0.362	0.159	0.185	0.217	0.097
Observations	423	173	423	183	80	183

Notes: For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Source: Authors' calculations

only 15 percent offer a correct one. The unconditional probability of a correct diagnosis was only 4 percent.

Private providers in the representative sample are more likely to offer a diagnosis but are not more likely to offer a correct one. The probability of offering a correct diagnosis is higher in the dual practice sample (15 percent versus 4 percent), which is not surprising since these providers are all trained MBBS doctors. Even among these doctors, however, there is no difference in the rate of correct diagnosis between their public and private practices. Overall, the summary statistics, our price regressions (seen later), and our field work suggest that pronouncing a correct diagnosis (or even just a diagnosis) is not seen by providers (and the market) as

TABLE 5—TREATMENT IN THE PUBLIC AND PRIVATE SECTORS (*Unstable Angina and Asthma Cases Only*)

	Representative sample					
	Correct treatment (1=Yes) (1)	Palliative treatment (1=Yes) (2)	Unnecessary treatment (1=Yes) (3)	Correct treatment only (1=Yes) (4)	Antibiotic (1=Yes) (5)	Number of medicines (dispensed and/ or prescribed) (Continuous variable) (6)
Panel A. SP and case fixed effects						
Is a private provider	0.052	−0.038	0.061	−0.008	0.016	0.972
	(0.045)	(0.056)	(0.072)	(0.023)	(0.062)	(0.279)
R^2	0.260	0.215	0.066	0.044	0.079	0.087
Observations	440	440	440	440	440	440
Mean of public	0.211	0.526	0.737	0.026	0.263	2.092
Mean of private	0.270	0.496	0.808	0.017	0.279	3.097
Mean of sample	0.266	0.498	0.802	0.018	0.278	3.021
Panel B. SP, case and market/district fixed effects						
Is a private provider	0.051	0.040	0.095	−0.020	0.086	0.894
	(0.051)	(0.059)	(0.079)	(0.027)	(0.084)	(0.264)
R^2	0.384	0.350	0.233	0.255	0.239	0.289
Observations	440	440	440	440	440	440
Panel C. SP, case and market/district fixed effects						
Is a private provider	0.101	0.060	0.066	−0.005	0.112	0.638
	(0.056)	(0.066)	(0.076)	(0.027)	(0.067)	(0.310)
Has MBBS	0.309	0.246	−0.132	0.106	0.267	−0.397
	(0.074)	(0.100)	(0.084)	(0.047)	(0.075)	(0.515)
Has some qualification	0.088	0.086	0.029	−0.001	0.099	−0.116
	(0.039)	(0.061)	(0.061)	(0.015)	(0.062)	(0.274)
Age of provider	−0.001	−0.001	−0.002	−0.000	−0.000	−0.012
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.011)
Gender of provider (1=male)	0.133	−0.118	−0.068	0.001	−0.029	−0.128
	(0.106)	(0.160)	(0.112)	(0.031)	(0.099)	(0.327)
Patient load during visit	−0.008	−0.017	0.007	−0.001	−0.008	0.009
	(0.008)	(0.012)	(0.006)	(0.001)	(0.006)	(0.026)
R^2	0.406	0.370	0.253	0.278	0.272	0.293
Observations	423	423	423	423	423	423

(*Continued*)

being essential in this setting. Note, however, that *pronouncing* a correct diagnosis is neither necessary nor sufficient for providing a correct or palliative treatment.[19]

D. *Treatment*

Table 5 reports on several outcomes related to the treatment offered, coded as discussed in Section IIC. The probability of receiving at least one correct treatment from a representative public provider was 21 percent. However, they offered

[19] Since the providers are usually much more educated than the typical patient, field interviews suggest that they often feel no need to explain themselves to the patients. Thus, providers may have an implicit diagnosis in their minds before they treat, but appear to feel no need to pronounce a diagnosis.

Table 5—Treatment in the Public and Private Sectors (*Unstable Angina and Asthma Cases Only*) (*Continued*)

	Dual practice sample					
	Correct treatment (1=Yes) (7)	Palliative treatment (1=Yes) (8)	Unnecessary treatment (1=Yes) (9)	Correct treatment only (1=Yes) (10)	Antibiotic (1=Yes) (11)	Number of medicines (dispensed and/or prescribed) (Continuous variable) (12)
Panel A. SP and case fixed effects						
Is a private provider	0.151	−0.126	−0.021	0.019	−0.141	0.002
	(0.061)	(0.057)	(0.054)	(0.026)	(0.067)	(0.200)
R^2	0.274	0.309	0.108	0.025	0.120	0.127
Observations	201	201	201	201	201	201
Mean of public	0.373	0.637	0.833	0.020	0.490	2.833
Mean of private	0.566	0.465	0.838	0.040	0.374	2.919
Mean of sample	0.468	0.552	0.836	0.030	0.433	2.876
Panel B. SP, case and market/district fixed effects						
Is a private provider	0.156	−0.127	−0.022	0.018	−0.139	−0.002
	(0.062)	(0.058)	(0.053)	(0.026)	(0.067)	(0.198)
R^2	0.299	0.315	0.167	0.039	0.135	0.155
Observations	201	201	201	201	201	201
Panel C. SP, case and market/district fixed effects						
Is a private provider	0.181	−0.106	−0.021	0.018	−0.122	−0.001
	(0.067)	(0.060)	(0.062)	(0.029)	(0.071)	(0.215)
Has MBBS						
Has some qualification						
Age of provider	−0.002	−0.007	0.001	−0.002	−0.001	−0.019
	(0.004)	(0.006)	(0.003)	(0.001)	(0.005)	(0.012)
Gender of provider (1=male)	0.049	0.097	0.111	0.007	0.152	0.286
	(0.115)	(0.092)	(0.091)	(0.033)	(0.103)	(0.330)
Patient load during visit	0.004	0.004	0.013	−0.004	−0.000	0.074
	(0.015)	(0.013)	(0.018)	(0.003)	(0.017)	(0.037)
R^2	0.279	0.318	0.180	0.053	0.164	0.180
Observations	183	183	183	183	183	183

Notes: For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample. In columns 6 and 12 the dependent variable is total number of medicines recommended to the patient (dispensed and/or prescribed). Medicines in the public sector are almost always dispensed at the clinic, whereas medicines in the private sector are both dispensed at the clinic and prescribed to be purchased elsewhere.

Source: Authors' calculations

non-indicated treatments at much higher rates, with a 53 percent probability of providing a palliative treatment and a 74 percent probability of providing an unnecessary treatment. Since the majority of providers provide unnecessary treatments, the probability of receiving only a correct treatment and nothing more is 2.6 percent. We can also examine two potential proxies for overtreatment: the rate of antibiotic prescriptions and the total number of medicines provided. Antibiotics were prescribed or dispensed in 26 percent of interactions (though they were not indicated for the asthma and angina cases), and an average of 2 medicines per interaction were dispensed.

In the representative sample, we do not find a significant difference between public and private providers on the probability of providing a correct, palliative, or unnecessary treatment; however, point estimates suggest that private providers have a higher probability of providing both correct and unnecessary treatments. Private providers in the representative sample also provide significantly more medicines (over 3 medicines on average, which is 50 percent greater than the public clinics).

In the dual practice sample, treatments provided in the private practice strictly dominate those provided in the public practice of the same doctor. The rate of correct treatment is 42 percent higher (16 percentage points on a base of 37 percent), the rate of providing a clinically non-indicated palliative treatment is 20 percent lower (12.7 percentage points on a base of 64 percent), and the rate of antibiotic provision is 28 percent lower (13.9 percentage points on a base of 49 percent) in the private practice relative to the public practice of the same doctor.

As discussed in Section IIC, the results reported here are based on treatments that were dispensed at the clinic as well as those that were prescribed. In public clinics, medicines were typically dispensed within the premises, and provided free.[20] In the private clinics, we observed both dispensing and prescribing behavior. While the fees charged included the medicines provided at the clinic, patients would have had to pay separately for prescribed medications. Since we do not observe the typical rate of patient adherence to prescription protocols, our results should be interpreted as referring to the quality of medical advice provided as opposed to the quality of realized health outcomes.

E. *Knowledge and Effort of Public and Private Providers*

There is a strong correlation between higher provider effort and the probability of giving a correct treatment (Figure 4). Nevertheless, the results in Tables 3 and 5 suggest that the higher effort exerted by private providers in the representative sample does not translate into better treatment outcomes. A natural explanation is that the representative private provider has a lower level of medical knowledge but compensates with higher effort, yielding comparable overall levels of treatment accuracy. To examine this possibility further, we use the "discrimination" parameter of each checklist item (as estimated by the IRT model; see online Appendix Table A.6), to classify individual items into terciles of low, medium, and high discrimination items.[21] Here, higher discrimination items are those that are more effective at distinguishing provider quality.

Online Appendix Table A.7 reports the same specifications as in Table 3 but compares public and private providers on checklist completion for different levels of item discrimination. All providers are less likely to complete high discrimination items on

[20] Medicines were provided free in 92 percent of interactions in the representative sample and 97 percent in the dual sample. Thus, for the most part, care in the public sector was free as it is meant to be.

[21] The classification of items into terciles of difficulty is done within each case, but the results are robust to classifying the items jointly across all cases as well. The terciles for each item are indicated in online Appendix Table A.6.

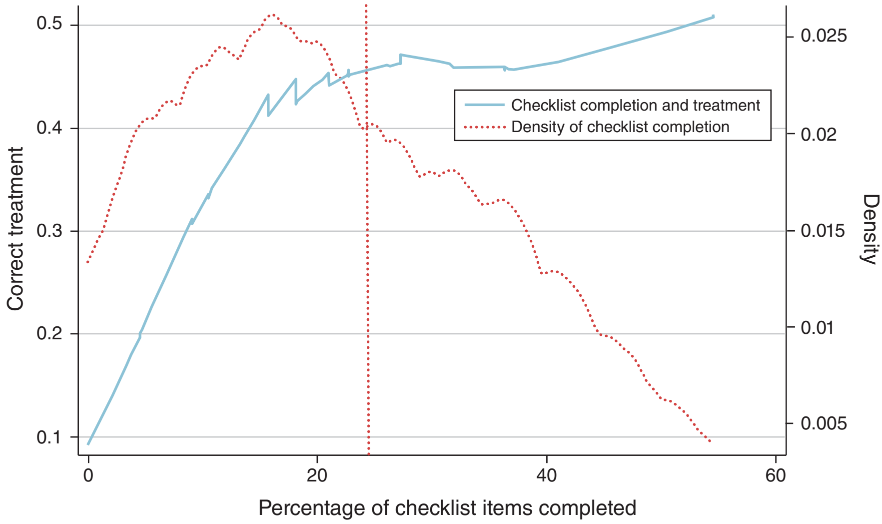

FIGURE 4. CHECKLIST COMPLETION AND TREATMENT

the checklist (consistent with low overall quality of care). In the representative sample, private providers complete 11 percentage points more of the low-discrimination checklist items but are no more likely to complete high-discrimination items. However, doctors in the dual sample are significantly more likely to complete both low and high-discrimination items in their private practice. These results suggest that while private providers do exert more effort, their lower knowledge leads to this effort being directed toward questions that are easy to ask and interpret, and may limit the marginal productivity of their effort. The results also highlight the importance of using the dual sample for holding provider knowledge and unobservable characteristics constant, and isolating the effect of market incentives on quality of care provided.

F. *Robustness of Checklist and Treatment Results*

Our main results pool data across cases to maximize power. For completeness, we also show the results from Tables 3–5 by case (online Appendix Table A.8). The superior performance of private providers on consultation length and checklist completion is seen in each of the three cases and in both the representative and the dual samples. Consistent with the overall results, private providers in the representative sample do not do better on diagnosis or treatment in any of the individual cases. In the dual sample, MBBS doctors were 14 percentage points more likely to correctly diagnose and 29 percentage points more likely to correctly treat the unstable angina (heart attack) case in their private practice relative to their public practices. In the asthma case, they are 13 percentage points more likely to offer a correct treatment (but this is not statistically significant given the smaller case-specific sample size).

We confirm that the results in Table 5 are robust to alternative definitions of correct treatment. Online Appendix Table A.9 shows the specific treatments offered by case, including referral frequency. Online Appendix Table A.10 shows that the results in Table 5 are robust to treating all referrals as a correct treatment. As discussed earlier, we include the dysentery case for the analysis of checklist completion but exclude it from the analysis of correct diagnosis and treatment because of the large (and differential) fraction of cases where the provider did not provide these and instead asked to see the child (see online Appendix Table A.9). Since checklist completion may also be censored in such cases, we also present the checklist completion results without the dysentery case and the results of Table 3 continue to hold (online Appendix Table A.11). We also show the core results with controls for clinic-level infrastructure and facilities (online Appendix Table A.12), and all the results continue to hold, suggesting that the results are not being driven by differences in facilities and infrastructure across public and private clinics. The final concern is that of differential completion rates of cases across public and private practices in the dual sample. We discuss this issue in detail in online Appendix D.1 and show that our estimates are likely to be a lower bound of the public private differences (online Appendix Tables A.13 and A.14).

IV. Results: Pricing and Cost Effectiveness

A. *Correlates of Prices Charged among Private Providers*

Private providers in this setting do not typically have fixed consulting fees that patients see or pay before entering the clinic. Rather, patients walk into the clinic and describe their symptoms. Pricing takes place at the end of the transaction with the provider asking the patient to pay a certain amount for the entire transaction including medicines dispensed. Our SPs followed this same protocol and paid the prices charged at the end of each interaction with a provider. We now examine the correlates of prices charged for completed transactions to understand what the market rewards in this setting.

Table 6 presents correlations between prices charged by private providers and our various metrics of health care quality in the representative sample, dual sample, and pooled sample. The odd columns present binary correlations, while the even columns present multiple regressions. The market rewards several measures of quality of care including time spent, checklist completion rates, and provision of a correct treatment (Table 6, columns 1, 3, and 5). On the other hand, there is no price premium for pronouncing a correct diagnosis and a price penalty for referrals; whether this penalty is optimal (without a penalty, every provider should just refer the patient) or reduces provider incentives to refer patients adequately is unclear. Finally, there is a price premium for dispensing medicines, but not for prescribing them.[22] The price charged is increasing in the total number of medicines dispensed, which may provide incentives for the provision of excessive medication.

Most of these patterns are repeated in the multiple regressions (Table 6, columns 2, 4, and 6). Note, however, that correct treatment is no longer rewarded in the multiple regressions. This is likely due to the high correlation between the

[22] Note that we cannot rule out the possibility that pharmacists provide doctors with a commission for prescriptions that they fill out, which would increase the incentives to over-treat as shown in China by Currie, Lin, and Meng (2014).

Table 6—Correlates of Price Charged (*Private Interactions*)

	Fees in Rs					
	Representative sample		Dual practice sample		Pooled sample	
	Binary regressions (1)	Multiple regression (2)	Binary regressions (3)	Multiple regression (4)	Binary regressions (5)	Multiple regression (6)
Time spent with SP (minutes)	1.763	0.771	2.498	2.017	1.502	0.805
	(0.454)	(0.475)	(0.587)	(0.679)	(0.361)	(0.390)
Percentage of checklist items	0.411	0.368	0.355	0.061	0.394	0.309
	(0.091)	(0.101)	(0.100)	(0.124)	(0.073)	(0.093)
Correct diagnosis (unconditional)	−3.749	−2.137	6.353	5.459	2.674	2.803
	(4.212)	(2.122)	(9.363)	(9.076)	(4.670)	(4.175)
Correct treatment	7.065	0.050	6.301	1.508	7.633	1.458
	(1.789)	(2.892)	(4.016)	(4.754)	(1.872)	(2.305)
Palliative treatment	8.036	5.581	11.748	7.798	8.124	6.252
	(2.056)	(2.036)	(4.344)	(4.663)	(1.811)	(1.863)
Unnecessary treatment	14.039	4.030	15.220	3.145	14.355	5.545
	(2.395)	(3.341)	(5.056)	(6.233)	(2.129)	(2.864)
Number of medicines dispensed	4.774	4.215	9.247	11.513	4.080	3.937
	(1.656)	(1.379)	(2.997)	(3.765)	(1.371)	(1.409)
Number of medicines prescribed	−0.202	−1.188	3.650	3.891	0.926	−1.020
	(1.129)	(0.881)	(1.845)	(2.672)	(0.861)	(1.067)
Referred/asked to see child	−19.161	−13.301	−10.082	−3.638	−16.857	−14.151
	(4.115)	(3.636)	(4.722)	(4.495)	(3.356)	(3.229)
Has MBBS	24.325	28.416			14.516	22.133
	(6.644)	(7.997)			(4.605)	(4.195)
Has some qualification	4.444	5.399			2.313	6.022
	(3.276)	(2.139)			(2.929)	(2.197)
Patient load during visit	0.736	0.441	0.276	0.029	0.503	0.149
	(0.665)	(0.333)	(0.863)	(0.876)	(0.602)	(0.510)
Age of provider	−0.150	−0.103	0.233	0.226	−0.095	−0.018
	(0.144)	(0.091)	(0.231)	(0.214)	(0.119)	(0.083)
Gender of provider (1=male)	−8.164	−4.923	−1.101	−3.713	−7.474	−3.098
	(3.497)	(4.969)	(4.845)	(5.460)	(2.918)	(4.069)
Constant		10.526		−11.589		3.386
		(6.561)		(12.095)		(5.913)
R^2		0.393		0.466		0.361
Observations		543		152		695
Mean price charged		27.327		33.125		28.699
SD		26.079		28.580		26.851

Notes: For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample and pooled sample, robust standard errors clustered at the location/market level are in parentheses. Observations are at the SP-provider interaction level. Interpretation of coefficients in "binary regressions" needs caution. Each coefficient represents a separate regression of prices on the row variable and SP, case and district fixed effects. Multiple regressions include SP, case and district fixed effects. The pooled sample (columns 5 and 6) combine the representative and dual practice samples.

Source: Authors' calculations

provision of a correct treatment and the checklist completion rate (Figure 4) and between correct treatment and the use of medicines. Thus the market appears to reward observable measures of quality such as time spent, checklist completion, and dispensing medicines (which are correlated with the provision of correct treatment),

Table 7—Wages in the Public Sector (*Public Observations Only*)

	log of monthly salary (pooled sample)		Desirability index (PHC/CHC sample)	
	Binary regressions (1)	Multiple regression (2)	Binary regressions (3)	Multiple regression (4)
Percentage of checklist items	0.002	−0.001	0.004	0.003
	(0.003)	(0.002)	(0.009)	(0.009)
Time spent with SP (minutes)	−0.051	−0.012	−0.061	−0.080
	(0.026)	(0.014)	(0.074)	(0.077)
Correct treatment	0.055	−0.090	−0.304	−0.132
	(0.066)	(0.048)	(0.237)	(0.202)
Has MBBS	1.055	1.283		
	(0.168)	(0.175)		
Has some qualification	−0.092	0.849		
	(0.367)	(0.300)		
Age of provider	0.012	0.019	0.052	0.062
	(0.006)	(0.006)	(0.019)	(0.024)
Gender of provider (1=male)	0.112	0.126	−0.530	−0.846
	(0.189)	(0.106)	(0.509)	(0.739)
Born in same district	−0.389	0.015	−0.180	0.101
	(0.147)	(0.081)	(0.449)	(0.432)
Is a dual provider	0.582	0.149	0.076	−0.135
	(0.136)	(0.086)	(0.402)	(0.527)
Constant		8.044		−1.470
		(0.316)		(1.198)
R^2		0.625		0.165
Observations		301		182

Notes: Robust standard errors are in parentheses. The pooled sample (columns 1 and 2) combine the representative and dual practice samples. The desirability index is constructed using principal component analysis of proximity to several amenities (paved road, bus stop, railway station, Internet, post office, and bank); availability of infrastructure (stethoscope, spyghamometer, torchlight, weighing scale, hand-washing facility, drinking water, staff toilet, patient toilet, fridge, sterilizers, electric connection, electric supply, power generator, telephone, computer, IV drip, cots/beds, disposable syringes); and PHC size (number of staff and number of patients). In binary regressions columns, each coefficient represents a separate regression of prices on the row variable, a constant and district fixed effects. Multiple regressions include district fixed effects.

Source: Authors' calculations

but patients do not appear to be able to discern whether they received the correct treatment conditioning on these observable measures.

The correlates of pricing observed in Table 6 point to both strengths and weaknesses of market-based incentives for health care provision. On one hand, there appear to be positive incentives for the provision of better quality care (including more effort and providing the correct treatment). On the other hand, the results are consistent with evidence from other settings, which show that markets for credence goods with asymmetric information between providers and customers often reward over-provision to the detriment of customer welfare. Overall, the results suggest that the market rewards providers who "do more," which is correlated with doing more "good" things as well as more "unnecessary" things.[23]

[23] Note that the results are robust to excluding observations where we were not able to identify the medicines provided and classify them as correct or not (see online Appendix Table A.15).

In sharp contrast to the market for private health care, the public sector rewards qualifications and age (experience), but there is no correlation between provider wages and any of our measures of quality including the time spent, checklist completion, or correct treatment (Table 7). Since public employees receive non-pecuniary rewards for better performance through more desirable job postings, we also present correlations between the desirability of a posting and measures of quality and again find that the only significant correlate of a better posting is age, suggesting that the public sector does not reward the quality of care provided by doctors with either more pay or with more desirable job postings.[24]

B. *Comparative Cost Effectiveness*

While health care in the public sector is free or nominally priced to the user, it is not cost-free to the tax payer. Online Appendix Table A.16 presents estimates of the cost per patient in the public sector, and calculates that the cost per patient interaction is around Rs 240. This is a conservative calculation because it uses only the wage cost in the public sector and does not include any cost of infrastructure, facilities, equipment, medicines, or administration. By contrast, the fees charged are the only source of revenue for private providers and hence will cover all operating costs. Thus, even though private providers charge higher consultation rates than public providers (as seen in Table 2), the per-consultation fee of Rs 51 charged by private providers is less than a fourth of the cost of a patient interaction in the public sector.[25]

V. Robustness

A. *Real Patients*

The use of SPs to measure quality of health care presents several advantages over the method of clinical observations. However, SPs are limited in the number and types of cases that can be presented. Further, we may worry that the SPs present "off equilibrium" situations in the market that do not extend to its general functioning. We therefore supplemented our data collection after completing the SP modules by conducting day-long clinical observations to code actual provider-patient interactions. We conducted these observations in both the representative and dual samples and in the latter observed a provider in both his/her private and public practices. While we cannot code the actual quality of care from these observations (since we do not observe underlying illnesses), we record several observable characteristics of each patient interaction based on over 1,000 interactions in both samples.

[24] These results are similar to those found in publicly-provided education in India and Pakistan, where teacher salaries increase with qualifications and seniority, but are not correlated with their effectiveness at raising test scores (Muralidharan 2013; Das and Bau 2014). Note also that our results add to a very limited evidence base (outside education) on the correlation between pay and productivity in the public sector, since worker-level productivity is typically not observed (see Muralidharan 2016 for a review of the evidence).

[25] Note that we assume that there is a comparable case mix for primary-health visits across public and private facilities, as is standard in comparative cost effectiveness analysis of this sort. This is also consistent with our data from observing real patients (see Section V below) where we observe considerable overlap in the symptoms presented across public and private clinics.

Table 8—Real Patients in the Public and Private Sectors

	Representative sample				
	Time spent (mins) (1)	Total questions (2)	Physical examination (3)	Dispensed/prescribed medicines (4)	Number of medicines (5)
Panel A. No patient or provider controls, and no fixed effects					
Is a private provider	1.456	0.799	0.371	−0.026	0.500
	(0.323)	(0.180)	(0.108)	(0.011)	(0.121)
R^2	0.054	0.030	0.103	0.003	0.017
Observations	1,137	1,137	1,133	1,138	1,138
Mean of public	2.378	2.994	0.473	0.994	2.319
Mean of private	3.833	3.793	0.844	0.968	2.819
Mean of sample	3.621	3.676	0.790	0.972	2.746
Number of public providers	29	29	29	29	29
Number of private providers	169	169	169	169	169
Panel B. No patient or provider controls, and market/district fixed effects					
Is a private provider	1.626	0.630	0.503	−0.016	0.674
	(0.490)	(0.170)	(0.112)	(0.014)	(0.167)
R^2	0.163	0.162	0.218	0.090	0.167
Observations	1,137	1,137	1,133	1,138	1,138
Panel C. Including patient and provider controls, and market/district fixed effects					
Is a private provider	1.190	0.654	0.522	0.009	0.602
	(0.313)	(0.246)	(0.085)	(0.014)	(0.145)
Has MBBS degree	−0.466	0.373	0.159	−0.025	−0.337
	(0.462)	(0.217)	(0.079)	(0.016)	(0.206)
Has some qualification	0.334	0.027	0.011	−0.035	−0.178
	(0.378)	(0.153)	(0.052)	(0.015)	(0.146)
Age of provider	−0.025	0.008	0.001	0.001	0.007
	(0.011)	(0.005)	(0.002)	(0.001)	(0.006)
Gender of provider (1=male)	−1.337	−0.744	0.009	0.008	−0.016
	(0.705)	(0.729)	(0.090)	(0.018)	(0.209)
Patient can easily dress themselves	−0.058	−0.307	0.009	−0.036	0.083
	(0.475)	(0.601)	(0.118)	(0.023)	(0.413)
Patient can easily do light household work	0.722	0.023	0.027	0.024	−0.191
	(0.215)	(0.216)	(0.036)	(0.020)	(0.184)
Patient can easily lift a 5kg bucket and walk for 100m	−0.433	0.179	−0.025	0.005	−0.233
	(0.266)	(0.181)	(0.032)	(0.010)	(0.112)
Patient can easily walk 200–300m	0.416	−0.304	0.081	−0.027	0.019
	(0.262)	(0.225)	(0.048)	(0.016)	(0.150)
R^2	0.303	0.331	0.348	0.119	0.306
Observations	835	835	833	835	835

(*Continued*)

Table 8 reports results from estimating equation (1) with data from real patient interactions. Private providers spend more time with patients, ask more questions, and are more likely to conduct a physical exam. They also give out more medicines on average. Results from the dual sample are also remarkably similar to those in Tables 3–5, with private providers still exhibiting higher effort but not providing more medicines. Thus, while our SPs present only three specific cases, our results from observing real interactions between patients and providers across the entire set

TABLE 8—REAL PATIENTS IN THE PUBLIC AND PRIVATE SECTORS (*Continued*)

	Dual sample				
	Time spent (mins) (6)	Total questions (7)	Physical examination (8)	Dispensed/ prescribed medicines (9)	Number of medicines (10)
Panel A. No patient or provider controls, and no fixed effects					
Is a private provider	1.894	1.154	0.143	−0.007	−0.017
	(0.570)	(0.319)	(0.064)	(0.010)	(0.134)
R^2	0.115	0.082	0.017	0.001	0.000
Observations	1,085	1,083	1,082	1,091	1,091
Mean of public	1.499	3.284	0.678	0.991	3.190
Mean of private	3.393	4.439	0.821	0.983	3.169
Mean of sample	1.899	3.527	0.708	0.989	3.185
Number of public providers	51	51	51	51	51
Number of private providers	40	40	41	41	41
Panel B. No patient or provider controls, and market/district fixed effects					
Is a private provider	1.910	1.155	0.154	−0.008	−0.013
	(0.561)	(0.315)	(0.061)	(0.009)	(0.139)
R^2	0.120	0.101	0.074	0.006	0.016
Observations	1,085	1,083	1,082	1,091	1,091
Panel C. Including patient and provider controls, and market/district fixed effects					
Is a private provider	1.570	0.561	0.072	−0.016	−0.016
	(0.547)	(0.255)	(0.060)	(0.012)	(0.166)
Has MBBS degree					
Has some qualification					
Age of provider	−0.003	−0.012	−0.002	−0.000	−0.017
	(0.015)	(0.009)	(0.003)	(0.000)	(0.008)
Gender of provider (1=male)	−0.495	−0.040	−0.103	0.007	0.034
	(0.369)	(0.374)	(0.099)	(0.013)	(0.257)
Patient can easily dress themselves	−0.679	−0.807	−0.115	−0.002	−0.319
	(0.516)	(0.315)	(0.075)	(0.031)	(0.203)
Patient can easily do light household work	−0.267	0.505	−0.009	0.026	0.057
	(0.318)	(0.182)	(0.047)	(0.015)	(0.139)
Patient can easily lift a 5kg bucket and walk for 100m	0.149	−0.253	−0.036	−0.017	0.103
	(0.197)	(0.170)	(0.050)	(0.008)	(0.118)
Patient can easily walk 200–300m	0.554	0.245	0.097	0.013	0.121
	(0.285)	(0.155)	(0.064)	(0.008)	(0.154)
R^2	0.168	0.356	0.197	0.042	0.155
Observations	809	808	807	810	810

Notes: For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. Observations are patient-provider interactions, and the sample has been limited to the SP sample. The regressions in panel C include controls for patients' characteristics and patients' presenting symptoms. Controls for patients' characteristics include: whether patient has no education, number of questions asked by patient, and patients' asset index. Controls for patients' presenting symptoms include: number of days patient has been sick, patients' ease in performing activities of daily living, and indicators for a number of presenting symptoms (fever, cold, diarrhea, weakness, injury, vomiting, dermatological problem, pregnancy, and pain). In columns 5 and 10 the dependent variable is total number of medicines recommended to the patient (dispensed and/or prescribed).

Source: Authors' calculations

of cases seen in a typical day are very similar to those from the SPs, suggesting that our SP-based results may be valid for a wider range of cases.

B. *Statistical Discrimination*

Another issue in interpreting our dual-sample results is the possibility that doctors expect to see different patients and cases across their public and private practices, and that the differences we observe do not reflect market incentives as much as statistical discrimination.

We address this concern in three ways. First, we note that the cases are both standard and ubiquitous in our setting, and it is therefore unlikely to be "off the equilibrium" path for a provider to see a patient with these symptoms in either public or private clinics. Second, the cases were chosen such that the optimal diagnosis effort and treatment protocol for an initial consultation for these symptoms should not vary by the affluence level of the patient or their ability to afford follow up treatments. Third, we conducted detailed exit interviews with a sample of patients from each clinic that we conducted physician observations in. While patients visiting private clinics are wealthier and have more education (in the dual sample), we find that there are not many differences on average in case characteristics across public and private clinics (see online Appendix Table A.17). In other words, for the majority of observable symptoms and patient characteristics, it is not the case that patients go exclusively to a public or private clinic, suggesting that our results are unlikely to be explained by statistical discrimination (see online Appendix D.2 for a more detailed discussion).

C. *Strategic Diversion of Effort in the Dual Sample*

A further issue in interpreting our dual-sample results is the possibility that doctors with private practices may deliberately underprovide effort in their free public practices to shift demand to their fee-for-service private practices (see Jayachandran 2014 for a similar example from education). While we cannot fully rule out this possibility, there is suggestive evidence against this. We compare public providers with and without a private practice and find that providers with a private practice are not any more likely to refer away an SP (online Appendix Table A.18). Providers with a dual practice do provide less effort in their public practices relative to those without a private practice, but the lack of any evidence of differences in referral rates suggest that these differences may reflect selection rather than strategic behavior, with more publicly conscientious doctors less likely to have a private practice.

The relevant policy question is whether doctors will start exerting more effort in their public practice if the option of private practice did not exist. But it is worth noting that private practice by public MBBS doctors was illegal in MP during the time of our study and that over 60 percent of providers still had a private practice, consistent with the idea that this is a low accountability environment.

D. *Alternative Comparisons in the Representative Sample*

Finally, our representative sample analysis compares the average public and private provider in a market, but it is not clear if the average is the correct metric for quality since patients can choose the best provider in the market. We therefore present an alternative comparison between the best public and best private provider in *each* market in online Appendix Table A.19 and find that our results are very similar to those in Tables 3–5.

VI. Theoretical Framework

The main contribution of this paper is in establishing key facts about the functioning of health care markets for primary care in settings of low-income and low state capacity for administering high-quality public health systems. In online Appendix C, we present a simple theoretical framework to help interpret the facts that we document. We model provider-patient interactions as comprising of two stages: consultation, and treatment; and characterize the optimal effort and treatment choices that a provider is likely to make with and without market incentives, and the effects of their choices on patient health outcomes.

We model the consultation stage as one of Bayesian learning for the provider. Patients present their initial symptoms to the provider, based on which he forms a prior distribution regarding the true ailment. Higher effort in the consultation stage yields a more precise posterior distribution of beliefs regarding the true ailment (provider effort and knowledge are complements). In the treatment stage, the choice of treatment is determined by a combination of the physician's desire to cure the patient (which is facilitated by a more precise diagnosis), and market incentives for overtreatment. Consistent with our empirical results on correlates of prices charged, we assume that providers in the private sector receive compensation for (observable) effort, as well as a piece rate for medicines dispensed.

The main insight from the model is that market incentives will typically lead to higher diagnostic effort (because this is observable to patients and rewarded with higher prices) but that the impact on health outcomes is ambiguous. In settings where the default effort level of doctors under low-powered incentives is reasonably high (as may be true in many high-income country settings), the costs of overtreatment under market incentives will likely exceed the benefits of higher diagnostic effort (which has diminishing returns). However, in settings where the default effort level is very low (as is true for public-sector health care providers in our setting), the benefits of higher diagnostic effort in the private sector may outweigh the costs of over-treatment under market incentives.

The framework may help shed light on why the quote from Arrow (1963) regarding the undesirability of market-based provision of health care may not fully apply to our setting, though it may be highly relevant to developed country contexts.[26] More generally, it highlights the need for caution in extrapolating insights obtained in high-income

[26]For instance, the US health care literature has paid considerable attention to the problem of over-treatment induced by fee-for-service compensation of providers (see Clemens and Gottlieb 2014 for an illustration). But the default level of provider effort in the United States is reasonably high even without market incentives (through a combination of higher-quality medical training and accreditation, peer monitoring of practice standards, and a functioning liability regime for malpractice), making over-treatment the more salient concern. Over-treatment may also be a more first-order concern when patients are insured and do not face the marginal financial costs of over-treatment, which is not true in our setting.

settings to developing country settings where there is much lower state capacity to implement policies as designed (Muralidharan, Niehaus, and Sukhtankar 2014).

VII. Discussion and Conclusion

We present the first set of results on the quality of public and privately provided primary health care in a low-income country that features a de facto unregulated private sector, using an audit methodology that accounts for differences in both provider and patient composition. The audit results based on three tracer conditions are similar to those using observations of real patients in the same clinics across all cases, suggesting that our audit results may be externally valid for primary care more broadly in this setting. While our results may not generalize beyond primary care to critical illnesses or tertiary care, it is worth highlighting that primary care is an essential first step in identifying the need for tertiary care and directing patients accordingly. In particular, our angina case represents a situation where self-triage was difficult and where the patient would not know if he needed to go the hospital.

Overall, our results suggest that patients in our setting have few good options for health care, public or private. Private sector providers exert higher effort but their effectiveness is ultimately limited by their low level of medical knowledge, as the majority do not have formal medical qualifications. Public sector clinics, though theoretically staffed by qualified providers, are characterized by lower provider effort. Posts are vacant and doctors frequently absent, so that even in a public sector clinic, the patient often sees a provider without formal training. Lower effort compared to the private sector offsets the benefits of more qualified providers in the public sector, and ultimately there is little difference in correct treatment or the overuse of incorrect medicines across a representative sample of public and private providers. Further, our estimates suggest that the public health care system in India spends at least four times more per patient interaction but does not deliver better outcomes than the private sector.[27]

Comparing the same provider in the public and private sector allows us to isolate the comparative effectiveness of market-based accountability in the private sector to administrative accountability in the public sector. The former performs better on all counts. Adherence to checklists and correct treatment rates are higher in the provider's private clinic with no differences in the extent of unnecessary treatments.

These results are consistent with the hedonic earnings-effort relationship in the private sector, which is absent in the public sector. Providers in the private sector earn more when they complete more of the medically necessary checklist and when they provide a correct treatment, showing that the market rewards certain key aspects of high quality. However, the market also rewards unnecessary treatments and patients frequently receive and pay for treatments that they do not need, a finding that mirrors consistent concerns regarding overtreatment in the literature on credence goods.

[27] These results mirror recent experimental evidence on primary education. Muralidharan and Sundararaman (2015) find that private schools in rural India deliver equal or superior learning outcomes than public schools, even though public schools spend three times more per student. Private school teachers are less qualified than public teachers, but exert much higher levels of effort. Thus, private providers in both primary health and education appear to make up for lower qualifications with higher effort, yielding outcomes no worse than those provided by the public sector—which have much higher costs per student/patient.

In contrast, we find no correlation between provider salaries and any measure of the quality of care they provide in the public sector.

Despite market incentives for over-treatment, one surprising result is that the rate of provision of unnecessary medication is equally high in the public clinics. Our theoretical framework provides a possible explanation for this result by showing that unnecessary treatments are not only driven by market incentives, but can also arise from low diagnostic effort. In our setting of low default effort in the public sector, the increase in diagnostic precision enabled by higher effort in the private sector may offset the incentives for over-treatment under market incentives, yielding no net difference in the provision of unnecessary treatment. Overall, our results suggest that in settings with low state capacity for high quality public service delivery, the effort advantage of the private sector may outweigh the credence good costs of privately provided health care.

Indian and global health policy debates have been hampered by a lack of empirical evidence on the quality of clinical interactions in the public and private sectors. Under the status quo, considerable attention has been focused on improving access and spending for publicly provided health care (Planning Commission of India 2013). Our results suggest that enthusiasm for the public sector as the principal source of primary care services in resource poor settings has to be tempered by the extent to which administrative accountability is enforced in the system and that poor incentives for effort may be a binding constraint to quality in the public system of health care delivery.

On the other hand, the marginal returns to better training and credentialing may be higher for private health care providers who have stronger incentives for exerting effort. Current policy thinking often points in the opposite direction, with a focus on hiring, training, and capacity building in the public sector on one hand (without much attention to their incentives for effort), and considerable resistance to training and providing legitimacy to unqualified private providers on the other (Reddy et al. 2011; Shiva Kumar et al. 2011; Planning Commission of India 2013).

This viewpoint is often justified by assuming that patients—particularly those who are poor and illiterate—make poor decisions regarding their health care. While certainly possible, a more nuanced understanding of patient behavior in low-income settings requires better empirical evidence on the actual quality of care obtained from different types of health care providers. Our paper presents some of the first evidence on this question, and expanding this methodology to other conditions and settings will allow for a richer understanding of the functioning of health care systems in settings with low resources and poor administrative capacity.

REFERENCES

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53 (5): 941–73.

Ayres, Ian, and Peter Siegelman. 1995. "Race and Gender Discrimination in Bargaining for a New Car." *American Economic Review* 85 (3): 304–21.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer, and Matthias Sutter. 2013. "What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods." *Review of Economic Studies* 80 (3): 876–91.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo. 2004. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review* 94 (2): 326–30.

Banerjee, Abhijit V., Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6 (2–3): 487–500.

Centre for Policy Research (CPR). 2011. *Mapping Medical Providers in Rural India: Four Key Trends*. New Delhi: Centre for Policy Research.

Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan, and F. H. Rogers. 2006. "Missing in Action: Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20 (1): 91–116.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review* 104 (4): 1320–49.

Coarasa, Jorge, Jishnu Das, and Elizabeth Gummerson. 2016. "Evaluating the Evidence on Public and Private Sector Quality of Care in Low and Middle Income Countries." Unpublished.

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing Antibiotic Abuse in China: An Experimental Audit Study." *Journal of Development Economics* 110: 39–51.

Das, Jishnu, and Natalie Bau. 2014. "The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers." Unpublished.

Das, Jishnu, and Jeffrey Hammer. 2005. "Which Doctor? Combining Vignettes and Item Response to Measure Clinical Competence." *Journal of Development Economics* 78 (2): 348–83.

Das, Jishnu, and Jeffrey Hammer. 2007. "Money for Nothing: The Dire Straits of Medical Practice in Delhi, India." *Journal of Development Economics* 83 (1): 1–36.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "In Urban and Rural India, a Standardized Patient Study Showed Low Levels of Provider Training and Huge Quality Gaps." *Health Affairs* 31 (12): 2774–84.

Das, Jishnu, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan. 2016. "Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.20151138.

Dhaliwal, Iqbal, and Reema Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 121: 1–21.

Dulleck, Uwe, and Rudolf Kerschbamer. 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods." *Journal of Economic Literature* 44 (1): 5–42.

Faxon, David, and Claude Lenfant. 2001. "Timing is Everything: Motivating Patients to Call 9-1-1 at Onset of Acute Myocardial Infarction." *Circulation* 104 (11): 1210–11.

Foster, Andrew D., and Mark R. Rosenzweig. 1994. "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health." *Review of Economics and Statistics* 76 (2): 213–27.

Gertler, Paul, and Christel Vermeersch. 2013. "Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes." National Bureau of Economic Research Working Paper 19046.

Goldberg, Pinelopi Koujianou. 1996. "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 104 (3): 622–54.

Grépin, Karen. 2014. "The Role of the Private Sector in Delivering Maternal and Child Health Services in Low-Income and Middle-Income Countries: An Observational, Longitudinal Analysis." *Lancet* 384 (S7).

Gruber, Jonathan, and Maria Owings. 1996. "Physician Financial Incentives and Cesarean Section Delivery." *RAND Journal of Economics* 27 (1): 99–123.

International Institute for Population Sciences and Macro International (IIPS). 2007. *National Family Health Survey (NFHS-3), 2005–06: India*. Mumbai: International Institute for Population Sciences.

Jayachandran, Seema. 2014. "Incentives to Teach Badly: After-School Tutoring in Developing Countries." *Journal of Development Economics* 108: 190–205.

La Forgia, Gerard, and Somil Nagpal. 2012. *Government-Sponsored Health Insurance in India: Are You Covered?* Washington, DC: World Bank.

Leonard, Kenneth, and Malkiory C. Masatu. 2005. "The Use of Direct Clinical Observation and Vignettes for Health Services Quality Evaluation in Developing Countries." *Social Science and Medicine* 61 (9): 1944–51.

Muralidharan, Karthik. 2013. "Priorities for Primary Education Policy in India's 12th Five Year Plan." *India Policy Forum* 2012–13 (9): 1–46.

Muralidharan, Karthik. 2016. "A New Approach to Public Sector Hiring in India for Improved Service Delivery." *India Policy Forum* 2015–16 (12).

Muralidharan, Karthik, Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, and F. Halsey Rogers. 2011. "Is There a Doctor in the House? Medical Worker Absence in India." http://econweb.ucsd.edu/~kamurali/research.html (accessed March 15, 2016).

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2014. "Building State Capacity: Evidence from Biometric Smartcards in India." National Bureau of Economic Research Working Paper 19999.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2015. "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India." *Quarterly Journal of Economics* 130 (3): 1011–66.

Pauly, Mark V. 1980. *Doctors and Their Workshops: Economic Models of Physician Behavior.* Cambridge, MA: National Bureau of Economic Research Monograph.

Planning Commission, Government of India. 2013. *Health, Nutrition and Family Welfare*. New Delhi: Planning Commission, Government of India. http://planningcommission.gov.in/plans/planrel/12thplan/pdf/12fyp_vol3.pdf.

Prendergast, Canice. 2003. "The Limits of Bureaucratic Efficiency." *Journal of Political Economy* 111 (5): 929–58.

Reddy, K. Srinath, Vikram Patel, Prabhat Jha, Vinod K. Paul, A. K. Shiva Kumar, and Lalit Dandona. 2011. "Towards Achievement of Universal Health Care in India by 2020: A Call to Action." *Lancet* 377 (9767): 760–68.

Rethans, Jan-Joost, Fred Sturmans, Cees van der Vleuten, and Pie Hobus. 1991. "Does Competence of General Practitioners Predict Their Performance? Comparison between Examination Setting and Actual Practice." *British Medical Journal* 303 (6814): 1377–80.

Rohde, Jon, and Hema Viswanathan. 1995. *The Rural Private Practitioner*. New Delhi: Oxford University Press.

Schneider, Henry S. 2012. "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair." *Journal of Industrial Economics* 60 (3): 406–33.

Selvaraj, Sakthivel, and Anup K. Karan. 2009. "Deepening Health Insecurity in India: Evidence from National Sample Surveys since 1980s." *Economic and Political Weekly* 44 (40): 55–60.

Shaban, Radwan Ali. 1987. "Testing between Competing Models of Sharecropping." *Journal of Political Economy* 95 (5): 893–920.

Shiva Kumar, A. K., Lincoln Chen, Mita Choudhury, Shiban Ganju, Vijay Mahajan, Amarjeet Sinha, and Abhijit Sen. 2011. "Financing Health Care for All: Challenges and Opportunities." *Lancet* 377 (9766): 668–79.

Wagstaff, Adam. 2013. "What Exactly is the Public-Private Mix in Health Care?" *Let's Talk Development* (blog), December 2. http://blogs.worldbank.org/developmenttalk/what-exactly-public-private-mix-health-care.

World Bank. 2003. *World Development Report 2004*. Washington, DC: World Bank.